

THE ARETAIC SIGNIFICANCE OF MORAL BELIEFS

Sara Copic

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfilment of the requirement for the degree of Master of Arts in the Department of Philosophy.

Chapel Hill
2016

Approved by:

Susan Wolf

Thomas E. Hill, Jr.

Russ Shafer-Landau

© 2016
Sara Copic
ALL RIGHTS RESERVED

ABSTRACT

Sara Copic: The Aretaic Significance of Moral Beliefs
(Under the direction of Susan Wolf)

I argue that a person's moral beliefs can contribute to or detract from the quality of that person's moral character. I argue that this happens when one's moral beliefs result from cognitive (as opposed to conative) features that contribute to or detract from the quality of one's moral character, and I argue that some cognitive features can do that. I also argue that moral beliefs contribute to a person's moral character when she self-consciously acts in accordance with those beliefs. Likewise, self-consciously flouting one's moral views detracts from a person's moral goodness. Since I maintain that conative states also play a part in constituting quality of moral character, my view is a hybrid view. This approach contrasts with the fully desire-based account of moral goodness that Arpaly and Schroeder present in *In Praise of Desire*. I present the desire-based account and show that a hybrid view has some explanatory advantages.

TABLE OF CONTENTS

Introduction	1
Spare Conativism.....	2
The Cognitive Side of Moral Character.....	12
The Moral Significance of Integrity.....	20
Explanatory Advantages.....	25
References.....	27

Introduction

I will argue that moral beliefs play an important role in contributing to and reflecting the quality of a person's moral character. While desires can and do contribute to and detract from the quality of a person's moral character, desires do not by themselves constitute the quality of a person's moral character or, in other words, a person's degree of moral goodness. This makes room for the view that beliefs can contribute to or detract from the quality of a person's moral character without deriving from desires. I will show that beliefs can do this in two ways.

A person's moral character is the set of features a person has that ground appropriate moral appraisal of that person, either considered as a diachronic unity or considered at a particular moment in time. For example, a person's moral character may contain a proper subset of the agent's desires, beliefs, and dispositions. Here, I will argue that beliefs deserve their place in this set in such a way that they don't have to derive from desires. I will not provide a strategy for comparing the extent to which beliefs and desires (either taken together or independently) contribute to the quality of a person's moral character in particular circumstances, nor will I present a way to calculate (so to speak) the extent to which a person has a good moral character, either over a fixed period of time or at a particular moment. I simply aim to show that a person's beliefs can, without deriving from that person's desires, contribute to the quality of the person's moral character.

Nomy Arpaly and Timothy Schroeder (2014) argue that good and ill will, praise- and blameworthiness, and moral virtue and vice are all ultimately explained by what desires we have and how we conceptualize those desires. In short, on their view, the quality of a person's moral character is ultimately constituted by our desires (that is, by a person's will) as long as they are conceptualized in the right way and beliefs are not part of this ultimate explanation. Belief only plays a derivative

role in reflecting the quality of person's moral character—a belief reflects the quality of a person's moral character only insofar as the person has that belief because of her intrinsic desires.¹

Moreover, while beliefs can *reflect* the moral quality of a person's will, they can never contribute to or (even partially) constitute the quality of a person's moral character. Actions partly based on moral beliefs reflect a person's moral character only insofar as those beliefs appropriately derive from the desires that constitute the character of a person's will. On their view, the quality of a person's desire set (that is, her *will*) fully grounds the quality of a person's moral character.

In what follows, I will present Arpaly and Schroeder's desire-based account of what it takes to have a good moral character. I will then argue for a hybrid account according to which both desires and beliefs can contribute to or detract from the quality of a person's moral character.

Spare Conativism

It is commonly thought that to have a good moral character is to have certain psychological features. Arpaly and Schroeder's view, which they call Spare Conativism (SC), is a view about which psychological features constitute what it is to be a good person or, in other words, what it is to have a good moral character and what kind of content and character these psychological features have. SC tells us that the only psychological features constitutive of a good moral character are certain kinds of desires. In fact, according to Arpaly and Schroeder, having a good moral character is fully constituted by having these kinds of desires; they form a complete explanatory base that grounds all of the notions connected with the moral appraisal of persons and they therefore specify the truth-conditions for judgments about the moral quality of a person's character. Since the basic units that constitute the quality of a person's moral character are desires (albeit with a particular content and

¹Instrumental and realizer desires do not contribute to goodness of will, and beliefs that reflect these desires therefore do not reflect goodness of will. See Arpaly and Schroeder (6-10, 87-88).

conceptualization), SC implies that the quality of a person's moral character just is the quality of the person's will or the person's desire set; the will is all there is to moral character.

According to SC, having good will consists in having the right intrinsic desires, correctly conceived, and having ill will consists in having the wrong intrinsic desires, correctly conceived.² The right intrinsic desires are desires for the right or the good, which is specified by the correct moral theory. Likewise, the wrong intrinsic desires are desires for the wrong or the bad, which is also specified by the correct moral theory. Interestingly, the way the person of good will must conceive of her desires in order to count as having good will is also specified by the correct moral theory, and likewise for the person of ill will. The person of good will must not only desire whatever is good or right, but must also conceive of her desires in a particular way in order to count as having good will.

One important clarification to note is that Arpaly and Schroeder's definitions of good and ill will do not trade merely on the distinction between desires for the good *de dicto* and desires *de re*. Arpaly and Schroeder are careful to emphasize this point with the help of an example (in their example, a phrase in all capital letters represents the conceptualization or mode of presentation of the desire):

If Lillian's intrinsic desire is for THE RIGHT while Shamissa's is to RESPECT PERSONS, and if Kantian normative theory is correct, then on our account the following things appear to be true: Lillian intrinsically desires the right or good *de dicto*, Shamissa intrinsically desires the right or good *de re*, and only Shamissa has so far been shown to have good will. However, both Lillian and Shamissa intrinsically desire the good *de re*. They both have the right itself, that very thing, as the object of their desires, as the referent of their respective concepts (165).

Why is it that Shamissa is the only one who has been shown to have good will, according to SC?

Both Lillian and Shamissa intrinsically desire the good *de re*, and hence both of them meet one necessary condition for having good will. But Shamissa has good will in virtue of conceptualizing

²This definition of good will is intended to be neutral between various first-order moral theories (Arpaly and Schroeder, 1).

this desire in the right way, namely, in the terms specified by the correct moral theory, which is supposed to be Kantian normative theory. To have good will, one must not only desire the good de re, but one's desire must be presented in terms of the concepts of the correct ethical theory.

Therefore Lillian, who desires the good de re and de dicto but does not conceptualize this desire in terms of respecting persons, does not have good will. While I do not find it plausible to say that Lillian does not have good will, I am going to set this aside. However, the distinction between de re and de dicto *beliefs* about the right will become important for my positive arguments for the moral significance of beliefs. We will get to the examples that are meant to motivate this view below.

It will be useful to list Arpaly and Schroeder's definitions in order to refer back to them later (162-163):

Complete good will is an intrinsic desire for [the whole of] the right or good, correctly conceptualized (i.e., presented via the correct sense, mode of presentation, narrow content, primary intension [...]).

Partial good will is an intrinsic desire for some part of the right or good, correctly conceptualized.

Complete ill will is an intrinsic desire for [the whole of] the wrong or bad, correctly conceptualized.

Partial ill will is an intrinsic desire for some part of the wrong or bad, correctly conceptualized.

Notice that, in the example above, and on the assumption of the truth of Kantian normative theory, Shamissa correctly conceptualizes her desire for the right and Lillian does not. This is because Lillian does not conceive of her desire as a desire to respect persons. Consider: a utilitarian might also desire THE RIGHT, de dicto and de re, without any further detail in his conceptualization of the desire. He desires the right, *whatever it is*, in those terms. But since the utilitarian doesn't conceive of her desire in terms of respect for persons, he does not count as having good will on this view. Likewise, Lillian only desires THE RIGHT, *whatever it is*. But this does not, according to Arpaly and Schroeder, suffice for good will.

Now, Arpaly and Schroeder also note that a person can lack good will without thereby having ill will, and that one can lack ill will without thereby having good will. So, we are also given the following definitions (163):

Moral indifference is a lack of good will. A person is more morally indifferent the less good will she has.

Reverse moral indifference is a lack of ill will. A person is more reverse morally indifferent the less ill will she has.

Lillian does not have ill will. This is because she fails to conceptualize her desire for the right correctly, namely, in Kantian terms. Rather, Lillian is merely morally indifferent since she fails to have good will. She does not have ill will because she does not desire the bad or the wrong.

One of SC's greatest achievements, according to Arpaly and Schroeder, is its explanatory power as compared to its parsimonious explanatory base. Good and ill will are defined in terms of intrinsic desires, correctly conceived, and all other morally significant features of moral character are defined in terms of good and ill will. First, acting for moral reasons is understood in terms of acting out of good will³. Second, Arpaly and Schroeder define virtue and vice in terms of good and ill will. To be virtuous is just to have significant good will and lack ill will; to be vicious is just to have significant ill will or significant moral indifference (202). Third, praiseworthiness and blameworthiness are also defined in terms of good and ill will. A person is praiseworthy for a right action just in case the action was rationalized by her good will or reverse moral indifference⁴, and a person is blameworthy for a wrong action just in case it was rationalized by her ill will or moral

³Generally, to think or act for a reason is for the event of one's thinking or acting to be caused by one's other attitudes in virtue of the fact that these attitudes rationalize the thought or action. To think or act for a moral reason is just a subset of thinking or acting for reasons. To think or act for a moral reason is just for one's thinking or acting to be caused by one's correctly conceptualized desires for the good or the right, that is, by one's good will. See Arpaly and Schroeder (62, 87-88).

⁴Lack of ill will, or reverse moral indifference, is an aversion to the wrong or the bad, correctly conceived. A person can be praiseworthy for doing something as a result of feeling an aversion towards doing harm, for example (163).

indifference.⁵ Indeed, it seems that all or most of the central notions connected to moral appraisal—acting for moral reasons, good and ill will, virtue and vice, praise- and blameworthiness—are defined in terms of correctly conceptualized intrinsic desires.

Moreover, the moral significance of beliefs is also explained in terms of intrinsic desires, correctly conceptualized.⁶ On SC, moral beliefs reflect the quality of a person's will just in case those beliefs manifest the person's intrinsic desires that constitute her good will, reverse moral indifference, ill will, or moral indifference. Insofar as her moral beliefs don't manifest the relevant intrinsic desires, correctly conceptualized, they do not reflect her moral character. This is because the quality of one's moral character is fully explained by reference to the referents and conceptualization of one's intrinsic desires.

The same explanation holds for other cognitive features besides beliefs. Arpaly and Schroeder admit that there is a cognitive side to virtue. For example, the virtuous person tends to notice morally significant details that others just don't see and she tends to be open-minded rather than prejudiced.⁷ Crucially, however, these cognitive features reveal the quality of a person's moral character only *if* they are manifestations of her good will or reverse moral indifference, that is, only if they are manifestations of good intrinsic desires (including aversions to the bad), correctly

⁵“Praiseworthiness: a person is praiseworthy for a right action A to the extent that A manifests an intrinsic desire (or desires) for the complete or partial right or good (correctly conceptualized) or an absence of intrinsic desires for the complete or partial wrong or bad (correctly conceptualized) through being rationalized by it (or them),” (170). Notice that this definition of praiseworthiness contains the definition of good will and the definition of reverse moral indifference, such that we can state it in this way: a person is praiseworthy for a right action A to the extent that A manifests complete or partial good will or reverse moral indifference through being rationalized by the relevant desire (or desires). For Arpaly and Schroeder, “a person is blameworthy for a wrong action A to the extent that A manifests an intrinsic desire (or desires) for the complete or partial wrong or bad (correctly conceptualized) or an absence of intrinsic desires for the complete or partial right or good (correctly conceptualized) through being rationalized by it (or them),” (170). Again, notice that the definition of blameworthiness contains the definitions of ill will and moral indifference in it.

⁶See Arpaly and Schroeder (163-166).

⁷See Arpaly and Schroeder (233-239, 241-245).

conceived. By the same token, a vicious person's cognitive features are vicious insofar as they manifest ill will or moral indifference.⁸ Arpaly and Schroeder see the cognitive side of virtue “as a content-efficacious but not rationalized manifestation of the good will that constitutes virtue” (226). This means that the moral goodness of the virtuous person's cognitive features is also ultimately explained by her desires, and likewise for the moral badness of the vicious person's cognitive features. I will return to this topic in detail in the next section.

Let us turn to some specific examples that will further explicate SC. I will return to these examples throughout the paper in order to show how my proposed account gives, as I'll argue, a better and more subtle treatment of them than the one offered by Arpaly and Schroeder. They use the following examples to support SC:

Brandon1: Brandon is an undergraduate whose moral views have been shaped by the writings of Ayn Rand. He believes that one's only moral duty is to act selfishly, and that actions taken for the sake of helping others are wrong and “pathetic”. Nevertheless, Brandon often acts selflessly. Brandon's selfless acts are in fact caused by the right intrinsic desires, correctly conceived. He simply doesn't notice how often he fails to act in accordance with the moral views he holds because helping others when they are in need just comes naturally to him. When his friend asks Brandon to help her move at the last minute, he heads over to help without giving it any thought. Suppose his only motivation for helping his friend is an intrinsic desire for whatever constitutes the right or the good correctly conceived. Brandon has no ulterior motives in helping his friend. Only later does he “realize he did the wrong thing” and berate himself for being “soft”.⁹

According to SC, Brandon1 does the right thing for the right reasons which, according to Arpaly and Schroeder, just amounts to being motivated by the right intrinsic desires, correctly conceived. The intuition we are supposed to get from this example is that Brandon1 is praiseworthy for his action and that he is “basically a good guy” who is unfortunately just bad at moral theorizing. Since being bad at theorizing—even moral theorizing—does not count against the quality of a person's

⁸“So the possessor of bad beliefs is vicious (and his actions based on the belief blameworthy), insofar as he is, either because his beliefs are manifestations of ill will or because they are manifestations of his deficiency of good will,” (236).

⁹I've altered the details of example slightly (see Arpaly and Schroeder, 177).

good will, Arpaly and Schroeder conclude that Brandon does not lack good will in virtue of his failure of moral reasoning (177).

Notice some elements that, according to Arpaly and Schroeder, do not or should not matter for our moral appraisal of Brandon1. First, the theory of which Brandon1 is convinced and the circumstances that led to his being convinced of it play no part in grounding the correct appraisal of his moral character. The quality of Brandon1's moral character is not impacted by how unreasonable, poorly supported, and out of sync objectivism is with the moral evidence we can expect Brandon to be exposed to and sensitive to—after all, Brandon1 is probably a college-educated young person living in the 1950s or later in the U.S. Of course, this is irrelevant if the quality of one's moral character is fully constituted by the quality of one's will. Second, failing to act in accordance with his moral theory does not detract from (or enhance) the quality of Brandon1's moral character, either; since his action is caused by a morally good desire, correctly conceived, flouting his own moral views does not detract from his moral character.

Brandon2: Brandon is an undergraduate whose moral views have been shaped by the writings of Ayn Rand. He believes that one's only moral duty is to act selfishly, and that actions taken for the sake of helping others are wrong and “pathetic”. When his friend asks Brandon to help her move at the last minute, Brandon consciously takes time to consider what's in it for him. He decides that helping his friend will benefit him more than if he refuses (free pizza and beer are almost certainly in his future if he helps, but he runs the risk of weakening a relationship that benefits him if he doesn't). Brandon judges that it is right to help his friend and does so (179).

In the case of Brandon2, Arpaly and Schroeder conclude that the agent does the right thing for the wrong reason, since he acts out of pure self-interest.¹⁰ Brandon2, they claim, does the right thing but he does it out of either partial ill will or partial moral indifference (more likely the latter). Therefore, insofar as we can extrapolate from this single action, Brandon2 is *not* “basically a good guy”.

Again, it's worth pointing out some features that are irrelevant according to SC, but that one

¹⁰For Arpaly and Schroeder, acting for reasons is understood in terms of acting on intrinsic desires (63-65). Very roughly, the idea seems to be that one acts for the right reasons just in case one's motives mirror the normative reasons

might reasonably think are relevant for the correct moral appraisal of Brandon². First, the theory of which Brandon¹ is convinced and the circumstances that led to his being convinced of it play no part in grounding the correct appraisal of his moral character. Second, Brandon² believes that he is acting rightly and he self-consciously acts in accordance with his moral views. That is, he has a moral conscience and self-consciously acts in accordance with it. Pre-theoretically, we judge it to be a good thing when a person's conscience constrains and even shapes a person's actions.¹¹ Therefore, it's worth noting that acting in accordance with his conscience (albeit a morally misguided conscience) does not, for Arpaly and Schroeder, enter into the explanatory base of the quality of a person's moral character (*unless* Brandon ² acts out the desire to act in accordance with conscience *and* this is a desire for the right or the good, correctly conceived). Finally, one way to understand what Brandon² is up to is not merely as acting on purely selfish desires, but also acting out of a desire to do what he believes is right. But according to Arpaly and Schroeder, it's doing the right things as a result of having the right intrinsic desires (correctly conceived) that counts, not doing the right thing because one believes it to be right and therefore worth doing.

A final example that Arpaly and Schroeder take to support SC:

Huckleberry Finn (Arpaly and Schroeder's interpretation): Huckleberry Finn befriends a runaway slave, Jim, and helps him escape to freedom. Huck judges it is wrong to do so and judges that it would be right to turn Jim in. But he simply cannot bring himself to turn in his friend. He desires to do the right thing, correctly conceived—we can suppose that he desires to promote his friend's well-being, under that description, and that promoting a person's well-being is part of the right or the good. Huck feels very guilty for knowingly doing “the wrong thing” though he in fact acts rightly (178).

Huckleberry Finn's case is purportedly like Brandon¹ in that Huck holds the wrong moral views but does the right thing for the right reasons, since he is motivated by the morally correct desire,

¹¹Though this is consistent with thinking that flouting one's conscience is better, all things considered, if a person has evil intentions.

correctly conceived.¹² Like Brandon1, Huck is also bad at moral theorizing and, according to Arpaly and Schroeder, our moral appraisal of Huck should not be affected by this fact. Notice that Huck self-consciously acts against his moral conscience. Yet on SC, this does not detract from the quality of his moral character, just as acting in accordance with his conscience does not contribute to the quality of Brandon2's moral character.

I agree with Arpaly and Schroeder on several points. First, I agree with the claim that Brandon1 is “basically a good guy”, but with a heavy emphasis on the “basically”. What I mean is that Brandon1 meets at least one of the intuitive conditions of having a partially good moral character: he has the right motives. This holds true for Huck Finn as well (though, as we will see, I will argue that there is a morally significant difference between Brandon1 and Huck). I also agree that Brandon2 has a morally worse character than Brandon1 and Huck; he doesn't even have the right motives, after all.

But I do not share all of their intuitions. I believe that Brandon1 and Brandon2 have partially bad moral characters in virtue of holding unreasonable moral theories. I deny the claim that one's moral views and moral belief-forming mechanisms reflect one's moral character just in case those beliefs and belief-forming processes causally derive from the right (or wrong) intrinsic desires, correctly conceived. Supposing that their moral views are not manifestations of morally bad desires, I will argue that Brandon1 and Brandon 2, though not Huck Finn, have a partly bad moral character in virtue of having the moral beliefs that they do.

I disagree with Arpaly and Schroeder on a second point. I believe that Brandon2 has a partly good moral character in virtue of self-consciously acting on the basis of the moral views he sincerely believes. Moreover, if Brandon1 *had* self-consciously flouted his moral views (he didn't,

¹²For another treatment of Huckleberry Finn's case, as well as the variety of conflicts between conscience and motives, see Bennett (1974).

since he only realized that he acted inconsistently with them after the fact), then he would—in virtue of having acted against his conscience—have a partly bad moral character. Huck Finn, because he acts against his conscience, also has a bad moral character in virtue of this fact but, as we will see, gets excused to some extent. All three of these characters may nevertheless be good people for the most part. My point is not that acting in accordance with one's moral views is necessary for being a good person at all, but that it is one of the features that contributes to being good or virtuous and its lack is a feature that can take away from a person's moral goodness.

In the next two sections, I will give positive arguments for these claims. I will challenge Arpaly and Schroeder's claim to have correctly characterized that in which good and ill will consists. I will argue that beliefs also play a role in reflecting in a non-derivative way the degree of a person's good will. It may sound odd to say that beliefs can, without deriving from desires, contribute to and reflect the moral quality of a person's *will*. One way to understand the will is as a set of desires or, more strongly, as the set of an agent's effective desires.¹³ I agree that we may use the term 'will' in this way, but this does not detract from my claim. My substantive claim is that the basic units in terms of which we should assess a person's moral character include not just desires, but also beliefs, in a non-derivative way.¹⁴ Thus, we should not think of a person's moral character as consisting of good and ill “will”. Rather, it would be better (and more revealing) to speak of a person's moral character, nature, or moral worth instead, and to say that a person's good and ill *will* are proper parts of her moral character or moral nature. This is opposed to Arpaly and Schroeder's suggestion that the quality of a person's will fully constitutes a person's moral character. Ultimately, I will argue that the notions of good and ill will, where 'will' is understood as a set of desires (intrinsic or not,

¹³See Frankfurt (1971).

¹⁴On my view, a person's moral commitments and judgments *may* derive from a person's desires, but the point is that they need not.

effective or not), provides too thin a basis for moral appraisal of persons.

Another way to put the point is this: Arpaly and Schroeder seem to claim that the quality of a person's moral character is ultimately a function of the quality of the person's will, understood as the set of a person's intrinsic desires, correctly conceived. By contrast, I claim that the quality of the will only partly contributes to the quality of a person's moral character. Being a good person also depends on the reasons a person ends up with her moral beliefs and on a disposition to act in accordance with one's moral commitments, views, or beliefs.

The Cognitive Side of Moral Character

According to Arpaly and Schroeder, virtue has a cognitive side (225). Virtuous people and vicious people not only differ with respect to their desires, but they also think in different ways.

Arpaly and Schroeder write:

Consider the kind person. She is more inclined to prevent suffering than most of us, and that is an obvious manifestation of her desire that people not suffer. She is also more inclined to feel bad about the suffering of others, and that too is an obvious manifestation of her good will. But the kind person is also going to have thoughts that are different from others' thoughts. The kind person will be more likely to notice suffering: *it captures her attention*. The kind person will be more likely to remember that the person she met three months ago was struggling to come to terms with her brother's death: *it sticks in her memory*. The kind person is more likely than others to see an angry man's outburst as a product of his feeling threatened and humiliated: *it seems plausible to her*. And in many other ways, the kind person's cognitive life will be different from the cognitive life of the less-kind person.

Consider also the open-minded person. She might have been raised with the idea that homosexuals are degenerates, but even so she is able to see that there is nothing degenerate about her homosexual labmate. She might dislike the fashion for wearing baseball caps backwards, but she is nonetheless able to see that several such backward-capped students in her class are very talented. The open-minded person is the opposite of the prejudiced person, in other words, in her cognitive life (225, my emphasis).

Recall that, on SC, being virtuous just is having significant good will and lack of ill will (i.e. reverse moral indifference).¹⁵ Arpaly and Schroeder explain why the virtuous person's cognitive life

¹⁵Here, as above, I understand the will to be the set of the agent's desires and I understand good will and moral indifference in the way defined above.

tends to be so different from the non-virtuous person's cognitive life as follows: the cognitive features typical of the virtuous person qua virtuous person are manifestations of her significant good will or reverse moral indifference in the sense that her good will or reverse moral indifference causes her to have the cognitive features she does.¹⁶ In short, the virtuous person has the cognitive features she does because of her good will or reverse moral indifference. Moreover, these cognitive features reflect the quality of the virtuous person's moral character only insofar as they causally derive from her good will. Therefore, having the set of cognitive features typical of a virtuous person cannot contribute to virtue or constitute what it is to be virtuous, even in part, since one's having this set of cognitive features is explained by one's already being virtuous, which just is having a significant degree of good will. If someone were to have the cognitive traits typical of the virtuous person but lack good will or if her significant degree of good will did not cause her to have these cognitive features, then they would not reflect the quality her moral character. (This kind of person is possible, on SC, though probably not likely to exist.)

I agree with Arpaly and Schroeder that the virtuous person has a cognitive life that reflects her virtuous character. Likewise, the vicious person has a cognitive life that reflects his vicious character. I disagree, however, that the explanation for the connection between being virtuous and having a particular set of cognitive traits is fully explained by the quality of the virtuous person's will. On my view, having a good moral character is partly constituted by having certain cognitive features that are not caused by a person's desires.

The virtuous person's cognitive features differ from the vicious person's cognitive features with respect to at least the following things: what captures her attention, what she tends to remember, what strikes her as plausible, which options seem open to her and which options seem

¹⁶“The strategy of this chapter is to explain the cognitive side of virtue as a content-efficacious but not rationalized manifestation of the good will that constitutes virtue,” (226).

closed, open-mindedness, epistemic humility as opposed to arrogance, what inferences she draws, which similarities she notices in morally similar situations and which features she judges to be irrelevant, and which features of people strike her as morally significant and which ones do not. I claim that these cognitive features (1) partly constitute what it is to be virtuous and (2) do not need to derive from morally good desires (correctly conceived or not) in order to constitute virtue. Both of these claims contradict SC, since, according to SC, cognitive features merely reflect or are merely evidence of a virtuous character (that is, they don't constitute what it is to be virtuous) and they must derive from a person's good will.

My argument is one by analogy. First, consider the cognitive features of a good musician. A good musician can notice when a chord is slightly out of tune and when the French horn comes in just slightly too early during rehearsal. She can hear how different musical lines fit together, and she can isolate them in her head as she listens to them play simultaneously. She can hear when the timbre of her cello is slightly too metallic as compared to the warm sound she imagines before she plays her solo. These cognitive achievements might at first require significant conscious effort, but may become involuntary over time. The good musician can hear things the rest of us can't while listening to the same symphony.

What she tends to remember, notice, and judge about the music she plays or listens to is what makes her a musical expert, and we do not need to posit any background desires in order for these cognitive features to explain why she is a good musician. All we need to explain the fact that she is a good musician is the fact that she has the cognitive features constitutive of being a good musician and her ability to play her instrument, neither of which stem from her desires.

Second, consider the good scientist. It is plausible that the good scientist has at least some of the following cognitive features: she has hunches that are borne out in experimental results (she tends to form good hypotheses), she creatively engineers solutions for setting up experiments, and

she can figure out why she did not get the experimental results that she expected. Just like the good musician, the good scientist possesses some cognitive features that make her a good scientist. We do not need to claim that these cognitive features causally derive from any desires in order for them to constitute what it is to be a good scientist.

Now, let's return to the observation that there is a cognitive side to virtue; the virtuous person has certain cognitive features that distinguish her from the vicious person, some of which are listed above. For the good musician and the good scientist, their cognitive features partly constitute what it is to be a good musician or scientist. Moreover, it does not seem plausible to claim that these cognitive features contribute to the quality of one's musical or scientific skills just in case they causally derive from desires. If we did not need to explain why the good musician or a good scientist by reference to cognitive features that derive from desires, then I do not see why we should say that the virtuous person's cognitive features (qua virtuous person) reflect the quality of her moral character only if they causally derive from desires. Perhaps being a good person is like being a good musician or scientist in this respect; to be virtuous is (among other things) to have certain cognitive features, and these cognitive features need not be caused by any moral desires in order to constitute what it is to be virtuous.

One might object to this argument as follows: we do not think the quality of a person's moral character is affected by merely cognitive features when these features seem to have nothing to do with morality. Therefore, the cognitive features that I listed above cannot constitute what it is to have a morally good character. For example, we don't normally think that reasoning poorly detracts from the quality of a person's character, so if a person reasons poorly about *moral* ideas it does not detract from her moral character. This is the intuition behind Brandon1, Brandon2, and Huck Finn, none of whom, according to Arpaly and Schroeder, have a bad moral character in virtue of reasoning poorly about morality. Another example the objector might cite is this: we don't fault

people for perceptual failures that result in false beliefs, so we should not fault a person for failing to be sensitive to moral evidence that results in false moral beliefs, either. In short, all cognitive failures are merely cognitive failures, *not* failures of character.

A second objection one might raise to my view comes from Kant:

It is impossible to think of anything at all in the world, or indeed even beyond it, that could be considered good without limitation except a good will. Understanding, wit, judgment and the like, whatever such *talents* of mind may be called, or courage, resolution, and perseverance in one's plans, as qualities of *temperament*, are undoubtedly good and desirable for many purposes, but they can also be extremely evil and harmful if the will which is to make use of these gifts of nature, and whose distinctive constitution is therefore called *character*, is not good (GW 4:393).

Here, Kant claims that “talents of mind”—which presumably include the cognitive features I claim, along with Arpaly and Schroeder, to be typical of the virtuous person—do not contribute to goodness of character because these cognitive features are only good when combined with an already good will. Those same cognitive features, when combined with an ill will, are actually bad-making features. The suggestion seems to be that a sadist, for example, can have the cognitive features typical of the virtuous person while his character is either not made better by that or is worsened by it. This suggests, like the first objection, that epistemic achievements and failures cannot contribute to goodness of moral character by themselves, if they do so at all.

But it does seem like our practices of moral appraisal sometimes involve judging people's characters for cognitive or epistemic failings. Consider arrogance. One way to understand arrogance is to see it as the attitude of taking oneself to be morally superior to others, and this would clearly be a morally criticizable feature that may be grounded in morally criticizable desires.¹⁷ But another plausible way to understand arrogance is as a cognitive failing that we find morally criticizable, although there are other cognitive failings that we think are morally neutral. This kind of arrogance plausibly involves the following features: weighing one's own opinion more heavily than the

¹⁷See Tiberius and Walker (1998).

opinions of others, a reluctance to change one's views in the face of disagreement with others who are experts on the relevant subject, significant reluctance to suspend judgment in the face of peer disagreement (or, perhaps, reluctance to conclude another person is one's epistemic peer), lack of sensitivity to evidence that one's views are false, and an exaggerated belief in one's practical competence. Sometimes a person may inherit these cognitive features as a result of having an antecedent desire to be right about some topic or as a result of wanting to be seen as an expert. But sometimes these cognitive features are not rooted in a person's desires. Someone may be arrogant and yet have the desires typical of a humble person (whatever these may be). Nevertheless, we often respond to people who exhibit what we might call *merely epistemic arrogance* with moral criticism.

Why is that? On the one hand, it may be because we psychoanalyze people a bit. We might judge (for example) that a person's reluctance to overturn his opinion in light of contradictory evidence from the experts reveals a morally bad desire to be superior to others. This explanation would appeal to Arpaly and Schroeder. But another reason we might judge the arrogant person to have a partly morally bad character is because he fails to take other people seriously, and taking others seriously is something we may have a moral obligation to do. However, taking others seriously constitutively involves adjusting one's views in light of expert testimony, for example. Taking other seriously may involve suspending judgment in light of peer disagreement or (more weakly) becoming less reluctant to suspend judgment in light of peer disagreement. So there is at least one plausible explanation for why we morally criticize people for epistemic failings that don't derive from morally bad desires—we criticize them for their moral failings even when those failings are or look to be just cognitive, not conative, failings.

This suggests that the relevance of cognitive features to the quality of a person's moral character does not require those features to be expressions of or be causally dependent on desires. Therefore, the quality of a person's character is not fully determined by the quality of the person's

will; quality of moral character depends at least partly on cognitive elements that are causally and explanatorily independent of desires. Since I agree that the quality of a person's moral character does partly depend on her desires, I will call my proposal *the hybrid account of moral character*.

Let us return to Brandon1, Brandon2, and Huckleberry Finn to see how my view applies to these cases. Since Brandon2 differs from Brandon1 with respect to his quality of will—after all, Brandon2 does not even have the right motives, whereas Brandon1 does—the hybrid account yields the result that Brandon2 does not have as good a moral character as Brandon1. For now, since this is the only respect in which they differ morally, I will treat them in the same way for the rest of this section and only speak of Brandon1. (Though in the next section, I argue that Brandon2 does have a good-making feature that Brandon1 lacks.)

Recall that, on SC, Brandon1 and Huckleberry Finn have the same quality of moral character since they both do the right thing for the right reasons—that is, they both do the right thing because they are motivated by an intrinsic desire for the right or the good, correctly conceived. Since both of them have false moral beliefs, both of them either have something wrong with their moral belief-forming mechanism, or with the environments in which those belief-forming processes operate, or both. According to the view I presented above, insofar as a person lacks the cognitive features that are typical of the virtuous person, those features detract from the quality of that person's moral character. The virtuous person, moreover, is sensitive to morally relevant evidence and reasons well morally, but Brandon1 and Huckleberry Finn both have a morally bad character insofar as they lack these cognitive features.

Plausibly, Brandon1 has bad moral belief-forming mechanisms but does not live in a misleading environment (i.e. an environment that is more likely to result in his having morally repugnant beliefs). If we fill in the details of Brandon1's case, we can see this clearly. Suppose he is living in the U.S. in 2016, that he is college-educated and that he is of average intelligence. Then it

seems that we could reasonably expect Brandon1 to have somewhat reliable moral belief-forming mechanisms; his moral sense can be expected to be sensitive to the right or the good. But for some reason, Brandon1 finds himself convinced of the view that his self-interest should be the only thing guiding his actions and that other people's well-being does not matter in its own right. Given that Brandon1's cultural environment is rich in evidence that would point a person with a reliable moral belief-forming mechanism in the right direction, we should conclude that something is wrong with his belief-forming processes. He is, moreover, missing just those cognitive elements constitutive of being virtuous. To the extent that he is missing those cognitive features, then, Brandon1 is not a good guy.

By contrast, Huck has both the misfortune of having a poor moral belief-forming mechanism and the misfortune of being immersed in an environment that presents him with morally misleading evidence. His environment is misleading because almost everyone he respects and trusts believes that slavery is permissible. Insofar as Huck's belief that turning Jim in is wrong results from both his poor belief-forming features and from his misleading environment, he is less criticizable for that belief than if he had formed it in an environment that was not misleading. Therefore, Huck Finn may have a better moral character than Brandon1, in spite of the fact that the moral view he holds is more repugnant than that of Brandon1. In fact, given the misleading evidence that impacts Huck's moral beliefs, he may not have as unreliable of a moral belief-forming process as Brandon1 does.¹⁸ After all, a virtuous person, in a systematically misleading environment, might end up with moral views that are way off the mark, just as a person with good vision will end up with perceptual beliefs that are way off the mark if she is immersed in a perceptually misleading environment. Nevertheless, Huck's moral belief-forming processes go wrong somewhere.

¹⁸Here, I am understanding the reliability counterfactually; for example, Huck's moral belief-forming mechanism is reliable to the extent that it would yield correct moral beliefs in an environment without any misleading evidence.

In this section, I hope to have shown that the quality of a person's moral character depends in part on cognitive features that do not causally derive from a person's desires. Insofar as a person's moral beliefs result from vicious cognitive elements, that person is morally criticizable for them in the sense that they reflect on the quality of her moral character. In the next section, I will argue that there is another way in which beliefs can contribute to the quality of a person's character without deriving from desires.

The Moral Significance of Integrity

In this section, I will try to show that being disposed to act self-consciously in accordance with one's moral views or one's conscience contributes positively to the quality of a person's moral character. By the same token, being disposed to self-consciously flout one's conscience detracts from the quality of one's moral character.¹⁹ Both of these claims are true, I will argue, no matter the content of one's moral beliefs.

To act with *integrity* is to act in accordance with what one sincerely and correctly takes to be one's deepest values and commitments—that is to say, to act with integrity is to self-consciously act in accordance with one's deep values; this includes moral as well as non-moral values and commitments.²⁰ Therefore, if a person acts with integrity she necessarily acts in accordance with an important part of her self-conception, or her understanding of who she is most fundamentally.

¹⁹I am not going to take a stand on what it is to have a disposition to behave in a certain way. Perhaps a disposition to act is just to have a standing desire for something. If that is so, then the moral value of acting in accordance with one's moral beliefs, insofar as it is morally valuable, will be intimately related to the desire to act in accordance with what one believes is right. I find this perfectly acceptable. Nonetheless, the content of this desire (from the point of view of the person whose desire it is) is *wanting to do what I believe is right*. So, depending on what a disposition is, the value of being disposed to act in accordance with one's conscience may turn out to be the value of a particular desire. Nevertheless, Arpaly and Schroeder seem to be unable to say that this desire contributes to quality of moral character since its content is not correctly conceptualized.

²⁰Here, I am stipulating this definition. I will go on to argue that this kind of integrity is good. For a substantive account of integrity and a summary of different views of the topic, see Calhoun (1995).

We have many ways to indicate that integrity matters to us, and that it sometimes acts as a constraint on what we do because it constrains what we think we ought to do: (1) “I couldn't live with myself if I did that (but you go right ahead)”, (2) “As a theist, I just can't do that,” (3) “Sometimes, you just have to do is what you think is best,” (4) “I have to stay true to myself.” Not only do we use statements like this to communicate that we hold ourselves to certain standards, but we also respect other people when they say things like this—even when we disagree about what they think they “have” to do. Of course, the meaning of (1)-(4) may change from context to context; for example (1) might be used to convey an inchoate form of moral relativism, and (4) may be utterly morally neutral in many contexts. But we also use these statements to express that acting in accordance with what we take to be our deep commitments to morality and to our personal projects matters to us. This bit of linguistic evidence motivates the view that we take acting with integrity to be a good thing.

Perhaps there are limits to how much we can respect someone who thinks that she has to do something life-threatening in order to stay true to herself; we might say, in cases like this, that she doesn't know what's good for her. That is, although staying true to oneself is something we take to be good, its significance may not outweigh prudential considerations. There are almost certainly limits to how much we would respect someone who, for example, thinks she has to stay true to her Nazi values by killing her children.²¹ Nevertheless, we often take the feeling conveyed by statements like “I have to do this” seriously, where the emphasis is not just on the “have to”, but also on the “I”. I think statements like (1)-(4) are morally significant, but not because sincerely believing one of them automatically makes it right to act on that basis (it doesn't). Rather, I think that acting on the basis of these beliefs (e.g., your belief that you couldn't live with yourself if you did that)

²¹Magda and Joseph Geobels murdered their five children the day they committed suicide on May 1, 1945, perhaps out of loyalty to Hitler.

contributes to a person's moral goodness or virtue to some extent.

That being said, I think we must leave open the possibility that a person can be truly evil although she has integrity. It may be the case that being a good person or being virtuous is a threshold concept. By this I mean that, in order to count as being a good person, someone has to meet a certain threshold—she has to exemplify a certain cluster of properties (or one of a number of clusters of properties), but that none of these properties alone is necessary or sufficient for meeting the threshold. But even if goodness of character is not a threshold concept, it's unlikely that our concept of extreme evil requires that the evil person has no good-making character features. So, my claim that integrity contributes to good moral character does not entail that villains who act in accordance with conscience cannot be villains after all.

What I have said so far entails that someone with evil moral commitments and desires is less evil in virtue of self-consciously acting on the belief that what she is doing is morally right. Acting with integrity is, I claim, something of a saving grace. However, this is not to say that someone like Huck Finn would have had a better moral character all things considered if he had acted with integrity. Recall that I am not giving a theory of how to weigh all of the factors that contribute to and detract from the quality of a person's moral character. It may be true that Huck has a better moral character than someone who would have followed her conscience. But someone else who follows her conscience and later realizes how wrong she was in doing so has a redeeming quality about her—namely, that she tried to morally do her best, though she failed miserably.

Let us turn to some examples that intuitively seem to support my claim that integrity contributes to goodness of character:

Anne and Julius are close friends and want to resolve a conflict between them. Anne asks a friend for advice, and in the course of explaining the problem tells her about a dark secret in Julius's past. Anne later tells Julius about the conversation. Julius feels betrayed. He asks Anne to lie to her friend in order to try to cover up the secret. Even though she believes lying is often morally permissible and realizes that lying would cause little harm to others and greatly benefit Julius, Anne refuses, citing the fact that being honest with those who are

closest to her is one of her deepest moral values.

Suppose that Anne in fact acts wrongly when she fails to lie to her friend. Even so, there still seems to be something good about her. This can be explained on the assumption that integrity is a good thing. Although Anne acted wrongly, she tried to do right according to her sincerely held moral commitments. We might say that Anne reasonably denies Julius' request to flout her deep moral values even though she was wrong to do so. Given that this is not (by hypothesis) a life-threatening situation for Julius (rather, the fact that the secret is out is something that he is uncomfortable with but not endangered by), we intuitively judge that Anne reasonably sticks with her conscience. This reasonableness that she instantiates cannot be explained by the fact that self-consciously acting in accordance with her moral view makes the action right, since we've supposed it is wrong. But integrity provides Anne with a *pro tanto* reason for refraining from lying. She did the wrong thing but for a right reason.

Anne's example, insofar as it shows that one has a *pro tanto* reason to act with integrity, suggests that we may not be able to separate moral psychology from normative ethics quite as much as Arpaly and Schroeder intend to do. Arpaly and Schroeder explicitly state a commitment to articulating a moral psychology that does not presuppose many, if any, *particular* first order ethical claims.²² They suggest that first order normative theories and moral psychology stand apart in that we can do one without doing the other. Examples like Anne's, however, seem to suggest otherwise.

Another example:

A father and son have been estranged for a long time as a result of the son's resentment.

²²“By being focused on our psychological relationship to morality, moral psychology does something different from other branches of moral theory. Normative moral theory tells us what it is right or wrong to do, and what states of affairs are morally good or bad. Meta-ethical theory tells us what kinds of things rightness, wrongness, goodness, and badness are in the first place. But moral psychology takes for granted that certain things are right or good, and other things wrong or bad. The job of moral psychology is to tell us how, in acting, feeling, and thinking, we end up related to the right and the good, or the wrong and the bad,” (1). Throughout the book, Arpaly and Schroeder try to show that their theory of moral psychology is compatible with plausible but inconsistent first order moral theories, which also suggests that they take on board a commitment to be neutral between these moral theories.

The son resents his father for being, in his view, far too strict with him. The father would use corporal punishment in the form of spanking the son when he was young in order to change his behavior. The father thought he was doing exactly what a good father should do. He would sincerely think to himself, “I don't *want* to do this, but I have to bring him up well.” Years later, when they speak openly about the past, the son learns that the father wanted to be a good dad and sincerely thought he was doing the right thing. Suppose that the father acted wrongly when he would frequently use corporal punishment.

Intuitively, it would be reasonable of the son to forgive his father because the father sincerely believed that he had to do those things in order to be a good father. He acted in accordance with his conscience. Insofar as one has the intuition that this would justify the son's forgiving his father, one should also conclude that the father had a partially good moral character in the past.

Arpaly and Schroeder are not able to say that the father has a partly good moral character as a result of following his conscience and doing the wrong thing; this is because his doing the wrong is either a desire for the wrong thing (*de re*), or a desire for the right thing (being a good father) wrongly conceived, depending on how we fill in the details. With respect to Anne, they would have to say that her desire not to lie to her friend is simply not the right intrinsic desire (since we supposed that Anne acted wrongly), and so she cannot have a good moral character in virtue of sticking with her conscience. Now Arpaly and Schroeder argue partly from examples, and partly from explanatory power; here, I am engaging with them by pointing out intuitions one might reasonably have in response to the examples I've given. Of course, no theory will be able to accommodate all of our intuitions. But it is worth pointing out that not only do these examples seem intuitive, but we also take statements like (1)-(4) as indicators of partial goodness of moral character.

If I am right about the value of acting with integrity, then Brandon₂, who does the wrong thing for the wrong reasons, still has a redeeming quality because he acts with integrity. Brandon₁ fails to act with integrity but does not self-consciously act in a way that flouts his commitments, since he realizes his mistake only later. Huck Finn fails to act with integrity in a stronger sense than

Brandon¹ because he does self-consciously flout his moral convictions. Insofar as one has the intuition that Brandon² is the worst of these three characters (and I do), however, one might think that integrity does not matter as much as having the right motives. Perhaps acting with integrity gives one's moral character a “boost”, but it cannot compensate for having a significant amount of ill will or moral indifference.

Explanatory Advantages

A hybrid view that accepts both the moral significance of desires and beliefs can explain the moral value of acting in accordance with one's commitments in urgent cases of peer disagreement and epistemic indeterminacy. “Urgent” here refers to the fact that one has to decide at some point what to do, and sometimes quite quickly. If deliberation has to stop—either because a person cannot figure out on his or her own what the morally right action will be in time, or because two people are disagreeing and the conflict cannot be resolved before it's too late—what should bring deliberation to a halt is not a coin flip, say, but one's own commitments. Integrity can serve as a tie-breaker in urgent situations. Moreover, if someone sticks with her conscience in such a situation, we would not fault her for it; we would not only think this was epistemically rational, but morally praiseworthy (provided it wasn't the result of thick-headedness or close-mindedness, or some other vice).

A hybrid view can also explain, much more naturally than Spare Conativism could, the moral significance of statements such as “Sometimes, you just have to do what you think is best” and “I couldn't live with myself if I did that.” Supposing that these statements are morally significant and supposing that they do not ground the rightness of the relevant action in question, Spare Conativism would have to paraphrase them into statements about intrinsic desires. But, it's not clear how such a paraphrase would go, since it would seem to make these statements lose their moral

significance. For example, “Sometimes, you just have to do what *you* think is best” could not be paraphrased into “Sometimes, you just have to do what *you* want,” since merely wanting something does not carry the same moral significance as the original statement. Stating “I was just doing what I thought was best” would excuse someone from blame in light of having acted wrongly (as we saw in the father/son example), but “I was just doing what I wanted” would never excuse someone from such blame. This suggests that the language of desires cannot fully capture the significance of acting on the basis of one's beliefs.

Likewise, “I couldn't live with myself if I did that” is not properly paraphrased into something like “I couldn't live with myself if I did something I didn't really want to do.” The paraphrases, then, would have to be more subtle than these. I am not saying it would be impossible to articulate them, but I doubt that we will be able to preserve the moral significance of the original sentences in most cases.

Of course, the statements in question (such as statements 1-4) might not be morally significant, after all. This may turn out to be an assumption we throw out as we get close to reaching reflective equilibrium. But our moral language is one piece of evidence that we should take seriously during inquiry.

REFERENCES

- Arpaly, Nomy and Timothy Schroeder. *In Praise of Desire*. Oxford University Press, Oxford: 2014. Print.
- Bennett, Jonathan. "The Conscience of Huckleberry Finn." *Philosophy* Vol. 49, No. 188 (1974). 123-134.
- Calhoun, Cheshire. "Standing for Something." *Journal of Philosophy* Vol. 92, No. 5 (1995). 235-260.
- Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* Vol. 68, No. 1 (1971). 5-20.
- Kant, Immanuel, and Mary Gregor. *Groundwork of the Metaphysics of Morals*. Cambridge: Cambridge University Press, 1998. Print.
- Tiberius, Valerie and John D. Walker. "Arrogance." *American Philosophical Quarterly* Vol. 34, No. 4 (1998). 379-390.
- Williams, Bernard. "Integrity." *Utilitarianism: For and Against*. Ed. J.J.C. Smart and Bernard Williams. New York: Cambridge University Press, 1973. Print: pp. 108-118.